Data Warehouse Project

First, we start by declaring the modeling process step by step

Modeling Process:

Step 1: in this step, we will define the objectives of our business.

1.Marketing process:

the marketing department need to gain a deep understanding of the behavior and preferences of the company's frequent flyers, which can inform future marketing and promotional strategies.

in this process we indicates a commitment to understanding customers on a detailed level, and tailoring marketing efforts to their specific needs and preferences.

2.Reservations process:

The involvement of the finance team in analyzing company profits further highlights the level of detail involved in the process. The analysis is not just focused on the customer experience but also takes into account the financial implications of the reservation process, such as revenue generated and costs incurred.

3.Customer care process:

The airline company's customer care process focus on providing detailed and specific customer service interactions before, during, and after the trip. The company's approach includes analyzing the type of interaction and problem severity, indicating a nuanced and detailed approach to customer service. The process is designed to collect and analyze data on customer feedback, complaints, and inquiries, with the aim of identifying patterns and areas for improvement. Overall, the high level of detail and

specificity in this process demonstrates the company's commitment to providing high-quality customer service and continuous improvement.

By collecting and analyzing data on customer interactions, the company can identify patterns and areas for improvement, which can lead to business enhancements. The level of detail and specificity in this process suggests that the company is committed to providing high-quality customer service and continuously improving their operations.

Step 2: in this step, we will ask questions we need our model to answer.

* What is the count of each gender on each flight?
* What is the count of customer complaints?
* What is the count of customer complaints and inquiries per flight?
* What is the most profitable flight?
* What is the average profit for each channel?
* Who is the customer that delivers the most profit?
* How many tickets are being reserved by class?

Step 3: in this step, we will define granularity for each business process:

1-Marketing:

In this process, the granularity includes every flights the frequent flyers take, what fare basis they pay, how often they upgrade, and how they earn and redeem their frequent flyer miles. The marketing department also wants to know if frequent flyers respond to special fare promotions, how long their overnight stays are, and what proportion of these frequent flyers have gold, platinum or titanium status.

2-Reservations:

In this process, the granularity includes every reservation process, which suggests a focus on the specific steps involved in making reservations, such as selecting travel dates, seat assignments, and payment options.

Additionally, the fact that reservations can take place through multiple channels, such as phone, website, and third-party platforms, indicates a recognition of the need to consider different customer preferences and behaviors.

3-Customer Services:

The process is designed to collect and analyze data on every customer feedback, complaints, and inquiries, with the aim of identifying patterns and areas for improvement.

Step 4: in this step, we will define the dimensions tables.

1-Aircraft

That dimension table holds data about all the airplanes that the company owns.

2-Airport

That dimension table holds data about the airports that our aircrafts use.

3-Flight

That dimension table is used to connect aircraft and airport dimensions to the marketing fact table.

4-Fare Basis

That dimension table provides information about the ticket class (economy, premium economy, business class, first class).

5-Customer

That dimension table holds data about the passengers.

6-Locations

That dimension table holds data about the countries and cities available to travel in the airport.

7-Employee

That dimension table holds data about employees that are working in the airport.

8-Customer Care

That dimension table contains information about customer satisfaction levels, inquiries, and complaints.

9-Flyer Miles

That dimension table holds data about the flyer miles classes that the company offers.

10-Promotion

That dimension table holds data about the promotions that the company offers to customers.

11-Upgrade

That dimension table holds data about all types of upgrades that the company offers to customers.

12-Date

That dimension table is used as a calendar for entities in the model.

13-Time

That dimension table is used as a calendar for entities in the model.

14-Channel

That dimension table holds information about the channels that the company supports for tickets reservation.

15-Payment Method

That dimension table holds data about the different methods of payment the company offers.

Step 5: in this step, we will define the Fact tables .

1-Marketing Fact

That fact table helps us with Flight tracking and collecting some useful data about it, tracking the marketing team performance, tracking customers that have the best response to an upgrade or promotion or use the flyer miles system, tracking the time when planes take off and land and tracking every employee in the airport with some information about him.

* Measurments in the table:
* Number of passengers
* Number of tickets
* Number of overnights

2-Reservations Fact:

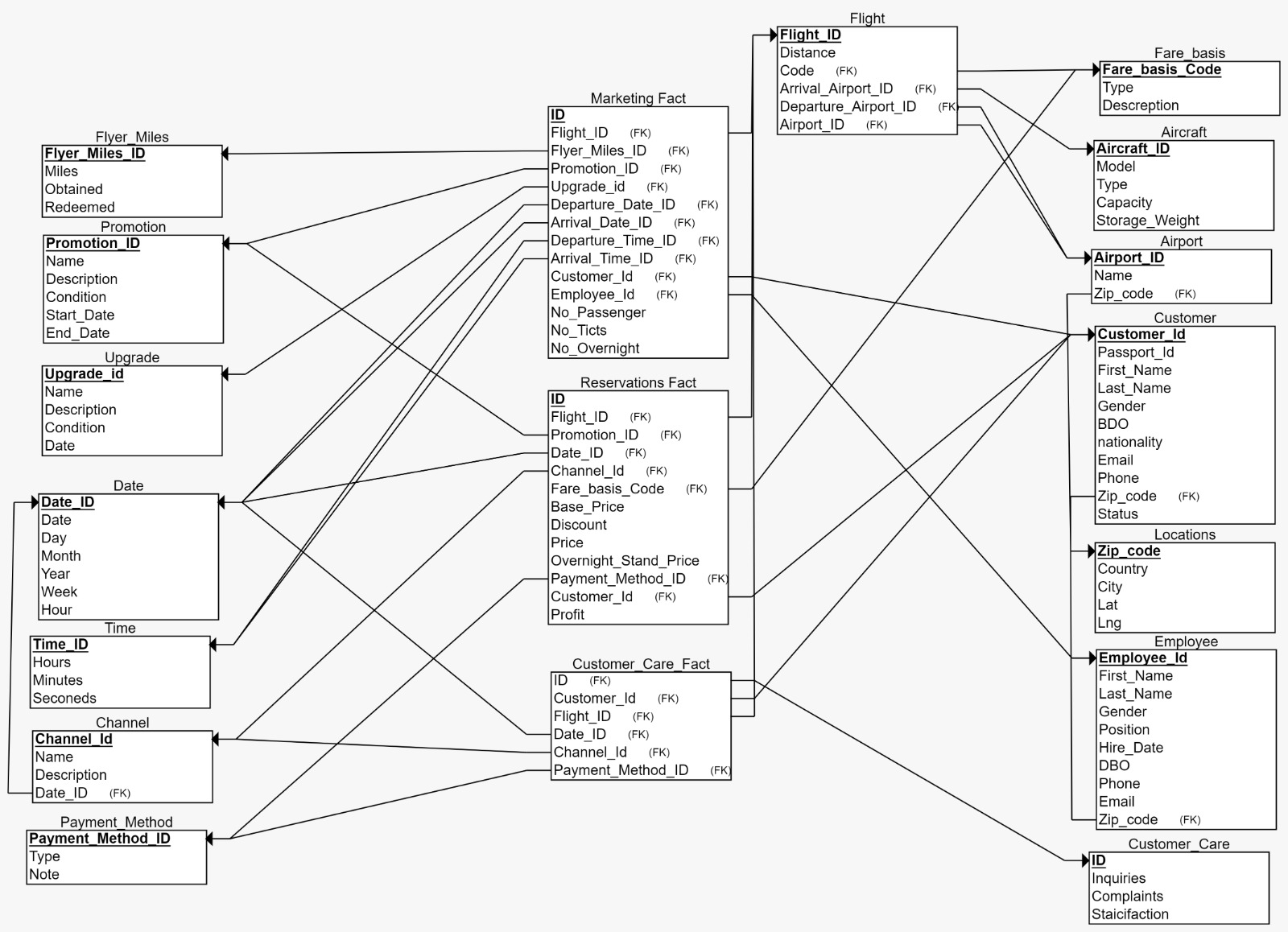
That fact table will help the financial team track customers that bring the most value, payment methods that bring the most money, flights that deliver the best profit, and reservation channels that are most rewarding.

* Measurments in the table:
* Base\_price
* Discount
* Price
* Overnight\_Stand\_Price
* Profit

3-Customer Care Fact:

That fact table helps us track customer satisfaction, inquiries, and complaints that will help us to improve our services.

Second, we will make the modeling schema.

Modeling Schema:

After making the model, now we will explain

Why did we choose snowflake modeling?

A snowflake schema is perfect for our design for some reasons:

* Reduces Data Redundancy: In a snowflake schema, dimensions are normalized into multiple tables, which can reduce the amount of data redundancy.

This means that each dimension table contains a subset of the data, and the data is not repeated in multiple tables. This reduces storage requirements, which can lead to improved performance and storage efficiency.

* Improves Query Performance: Snowflake schema allows for efficient querying of data. By normalizing dimensions into separate tables, it reduces the number of redundant data points that need to be queried. This means that queries can be executed faster and more efficiently, leading to better performance.
* Easier Maintenance: The snowflake schema is easier to maintain than a star schema because it reduces the amount of data redundancy. When changes are made to the schema, it only needs to be updated in one place, which reduces the risk of errors and makes maintenance easier.

This script we created to generate the fact tables:

SET SERVEROUTPUT ON SIZE UNLIMITED

declare

max\_id number;

rand\_start number := round(***dbms\_random.value***(1,600));

rand\_end number := round(***dbms\_random.value***(1,600));

x number;

cursor Marketing\_fact\_cursor is

select \* from p\_Marketing\_fact;

begin

for fact\_row in Marketing\_fact\_cursor

loop

select max(ID) into max\_id from p\_Marketing\_fact;

max\_id := max\_id+1;

if rand\_start > rand\_end then

x := rand\_start;

rand\_start := rand\_end;

rand\_end := x;

end if;

if (rand\_end - rand\_start) > 100 THEN

rand\_end := round(rand\_end / 10);

if rand\_start > rand\_end then

x := rand\_start;

rand\_start := rand\_end;

rand\_end := x;

end if;

end if;

FOR i in rand\_start..(rand\_end - 1)

LOOP

max\_id := max\_id+1;

if i != fact\_row.CUSTOMER\_ID then

insert into p\_Marketing\_fact (id, CUSTOMER\_ID, EMPLOYEE\_ID, FLYER\_MILES\_ID, PROMOTION\_ID, UPGRADE\_ID, ARRIVAL\_DATE\_ID, FLIGHT\_ID, NO\_OVERNIGHT, DEPARTURE\_DATE\_ID)

values(max\_id , i ,fact\_row.EMPLOYEE\_ID, fact\_row.FLYER\_MILES\_ID, fact\_row.PROMOTION\_ID, fact\_row.UPGRADE\_ID, fact\_row.ARRIVAL\_DATE\_ID, fact\_row.FLIGHT\_ID, fact\_row.NO\_OVERNIGHT , fact\_row.DEPARTURE\_DATE\_ID);

end if;

END LOOP;

end loop;